1 Appendix 1: Statistical Model Development and Implementation

The state-space model joins statistical models of observation methods with a process model portraying the underlying mechanisms of movement and population growth. The model can be conceptualized as a time series of unobserved true states (e.g. latent states) with the current state directly affecting the state at the next time. A second time series runs in parallel which are the observations of these true states made with error. Hierarchical bayesian techniques provide a framework for factoring such highly dimensional problems into lower dimensional ones. These techniques decompose a problem into the underlying process, data, and parameters, and identify uncertainty associated with each component.

Process Model

The spatial distributions of bison were estimated monthly during July through peak migration during 1990-2012. The column vector $\mathbf{z_{t,j}}$ was used to represent the number of bison in each wintering area during month t and year j. The matrix $\mathbf{A}$ defined transitions of bison between wintering areas during each month, where $\mathbf{z_{t,j}} \sim gamma(\beta, \mathbf{A}\mathbf{z_{t-1,j}}\beta)$ and

$$z_{t,j} = \begin{bmatrix} Central\, herd\, in\, Hayden\, Valley \\ Central\, herd\, in\, Firehole\, River\, Drainage \\ Central\, herd\, in\, Gibbbon\, River\, Drainage \\ Central\, herd\, in\, Hebgen\, Lake\, Basin \\ Central\, herd\, in\, Blacktail\, Deer\, Plateau \\ Central\, herd\, in\, Gardiner\, Basin \\ Northern\, herd\, in\, Lamar\, Valley \\ Northern\, herd\, in\, Lower\, Yellowstone\, River\, Drainage \\ Northern\, herd\, in\, Blacktail\, Deer\, Plateau \\ Northern\, herd\, in\, Gardiner\, Basin \end{bmatrix}$$

$$
\mathbf{A} = \phi \begin{bmatrix}
(1-\gamma_1)(1-\gamma_5) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\gamma_1 & (1-\gamma_2) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\gamma_5 & \gamma_2 & (1-\gamma_3)(1-\gamma_4) & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \gamma_4 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \gamma_3 & 0 & (1-\gamma_6) & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \gamma_6 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & (1-\gamma_7) & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \gamma_7 & (1-\gamma_8) & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \gamma_8 & (1-\gamma_6) & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \gamma_6 & 1
\end{bmatrix}
$$

Movement probabilities ($\boldsymbol{\gamma_i}$) were related to time since the establishment of snow pack, annual snow pack magnitude, annual herd size, and year of study. Time since the establishment of snow pack was treated as a second order polynomial to allow for movement probabilities to reach annual minima which corresponded optimal months of the year for bison to occupy wintering areas. The parameter ϕ represented monthly survival during the migration period and was not related to age and sex.

A similar approach was used to estimate numbers of bison in each of ten demographic stages during July 2003-2012. The column vector $\mathbf{w_j}$ was created to represent the total number of bison in each stage during year j, where

$$
z_j = \begin{bmatrix}
Central\, herd\, 2mo.either sex \\
Central\, herd\, 13-15\, mo.\, female \\
Central\, herd\, >26\, mo.\, female \\
Central\, herd\, 13-15\, mo.\, male \\
Central\, herd\, >26\, mo.\, male \\
Northern\, herd\, 2mo.either sex \\
Northern\, herd\, 13-15\, mo.\, female \\
Northern\, herd\, >26\, mo.\, female \\
Northern\, herd\, 13-15\, mo.\, male \\
Northern\, herd\, >26\, mo.\, male
\end{bmatrix}
$$

The matrix $\mathbf{B}$ represented annual transitions between demographic stages. Parameters of the transition matrix were $\psi_{c,a}$ = central herd adult annual survival, $\psi_{c,c}$ = central herd calf survival, $\psi_{n,a}$ = northern herd adult annual survival, $\psi_{n,c}$ = northern herd calf survival, τ_c = central herd births, τ_n

= northern herd births, π = fetal sex, and ϵ = emigration of central herd members to the northern herd, where $\mathbf{B}$ =

$$\begin{bmatrix}
0 & 0 & \psi_{c,a}\tau_c(1-\epsilon) & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\psi_{c,c}(1-\epsilon)(1-\pi) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \psi_{c,a}(1-\epsilon) & \psi_{c,a}(1-\epsilon) & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\psi_{c,c}\epsilon(1-\pi) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \psi_{c,a}(1-\epsilon) & \psi_{c,a}(1-\epsilon) & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \psi_{c,a}\tau_c\epsilon & 0 & 0 & 0 & 0 & \psi_{n,a}\tau_n & 0 & 0 \\
\psi_{c,c}\epsilon(1-\pi) & 0 & 0 & 0 & 0 & \psi_{n,c}(1-\pi) & 0 & 0 & 0 & 0 \\
0 & \psi_{c,a}\epsilon & \psi_{c,a}\epsilon & 0 & 0 & 0 & \psi n, a & \psi n, a & 0 & 0 \\
\psi_{c,c}\epsilon\pi & 0 & 0 & 0 & 0 & \psi_{n,c}\pi & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \psi_{c,a}\epsilon & \psi_{c,a}\epsilon & 0 & 0 & 0 & \psi_{n,a} & \psi_{n,a}
\end{bmatrix}$$

Removals to the population occurred during November-April of each year through gather-and-consignment or hunter harvest. The vector H_j was created to represent the number of animals removed from each demographic stage during year j. The population was offset by these removals prior to growth, such that $\mathbf{w_j} \sim LogNorm(\mathbf{B}(\mathbf{w_{j-1}} - \mathbf{H_j}), \sigma)$.

Data Models

Two measures of monthly distributions were recorded, including aerial counts and GPS recorded locations of individual adult female bison. Aerial counts did not occur exactly one month a part and we needed to adjust our process model such that the time step was variable. We aligned model updates with counts. When a count did not occur during a month, we updated the model on the 15th day. Transition probabilities were adjusted to the appropriate time step by approximating survival as $\phi_t = e^{log(\phi)\Delta t}$ and movements as $\gamma_t = 1 - e^{log(1-\gamma)\Delta t}$. Counts were related to numbers of bison in wintering areas by treating them as Poisson-gamma variables where $\mathbf{y_{t,j}} \sim Poisson(\lambda_{\mathbf{t,j}})$ and $\lambda_{\mathbf{t,j}} \sim gamma(\alpha, \mathbf{z_{t,j}}\alpha)$. GPS recorded locations were related to numbers of bison in wintering areas by treating monthly relocations of all animals as multinomial random variable of size equal to the total number of bison fit with telemetry devices and multinomial probabilities were proportions of bison in wintering areas predicted by the process model.

Three measures of population demographics were recorded, including replicate aerial counts of total abundance during the breeding period, annual peak-calving aerial counts of newborns, and replicate ground surveys of numbers of animals in age and sex categories. Replicate breeding range counts were related to herd abundance by treating each count as a Normal random variable, where $y_j \sim N(\mathbf{w_j}, \sigma_{obs})$ and σ_{obs} was a parameter reflecting counting error. During calving aerial counts, older bison were differentiated from newborns. Total numbers of calves were treated as a random variable with binomial probability equal to the proportion of calves in the population predicted by the process model. Bison were located in two group categories during ground classification with animals either found in mixed gender groups or bull-only groups. Aerial counts coincided ground classification which enabled the estimation of the proportion of animals in each group type. We defined p_m as the binomial probability that bison were found in mixed gender groups. Ground classifications were related to process model predictions of herd age and sex structure by treating these ground observations as multinomial random variables of sizes equal to total numbers of animals classified in mixed gender groups and multinomial probabilities equal to $\mathbf{w_j} p_m$.

Total annual removals were known, but data on removals with regards to each demographic stage were imperfect. Removals near the western park boundary exclusively occurred to central herd animals. Therefore, we estimated annual multinomial probabilities of removal to central herd demographic stages ($\mathbf{r_j}$) and then multiplied total known removals (R_j) by these multinomial probabilities to estimate all removals for each year ($\mathbf{H_j}$). However, both central and northern herd animals were removed near the northern park boundary. So, the movement process model was used to first estimate the binomial probability that animals in the Gardiner Basin wintering area were from the northern herd. This binomial probability was then multiplied by multinomial probabilities of removals to each age and gender stage and total known removals, e.g. $\mathbf{H_j} = \mathbf{r_j} * p_{north} * R_j$ and p_{north} was the binomial probability of animals in the Gardiner basin belonging to the northern herd.

Parameter Models

Diffuse parameter models were used for all parameters related to movements. Therefore, all parameters with real support were treated as N(0,1000) random variables (e.g. $\gamma_{\mathbf{i}}$). Parameters with positive support were treated as N(0,1000) random variables on the log scale and parameters with support bound between 0 and 1 were treated as Beta(1,1) random variables.

Informative parameter models were used for parameters related to demographics when existing data from mark recapture studies were available. Otherwise, diffuse parameter models were chosen.

Table 1: Informative parameter models based on mark-recapture studies completed during 1995-2012 in Yellowstone National Park. These prior distributions were estimated using random effect models fit to annual observations of adult female bison.

Parameter	symbol	mean	standard deviation
Central Herd adult survival	$\psi_{c,a}$	0.86	0.19
Central Herd births	τ_c	0.74	0.15
Central to Northern emigration	ϵ	0.04	0.06
Northern Herd adult survival	$\psi_{n,a}$	0.94	0.12
Northern Herd births	τ_n	0.78	0.12

Implementation

Marginal posterior distributions of latent states and parameters were estimated using Markov Chain Monte Carlo (MCMC) and Metropolis-Hastings decision criteria. Each MCMC algorithm was run for 100,000 iterations and the first 25,000 iterations were discarded to allow for burn-in. Three MCMC chains were constructed for each unknown marginal posterior distribution and we confirmed convergence using the Gelman and Rubin test statistic by identifying a potential scale reduction factor <1.02 for each variable. Trace plots of marginal posterior distributions were inspected to ensure reasonable exploration of the parameter space and a Metropolis-Hastings acceptance rate near 0.40. All statistical analyses were completed using program R.